



[12] 发明专利说明书

专利号 ZL 200410011828.3

[45] 授权公告日 2007 年 8 月 1 日

[11] 授权公告号 CN 1330124C

[22] 申请日 2004.9.22

[21] 申请号 200410011828.3

[30] 优先权

[32] 2003.12.1 [33] US [31] 10/726,002

[73] 专利权人 国际商业机器公司

地址 美国纽约

[72] 发明人 刘震 乔治·V·波佩斯库

萨姆比特·萨胡

[56] 参考文献

WO97/19411A1 1997.5.29

CN1229219A 1999.9.22

WO03/043286A2 2003.5.22

CN1239252 A 1999.12.22

WO99/63710A1 1999.12.9

CN1117616A 1996.2.28

审查员 刘玲斐

[74] 专利代理机构 中国国际贸易促进委员会专利
商标事务所

代理人 康建峰

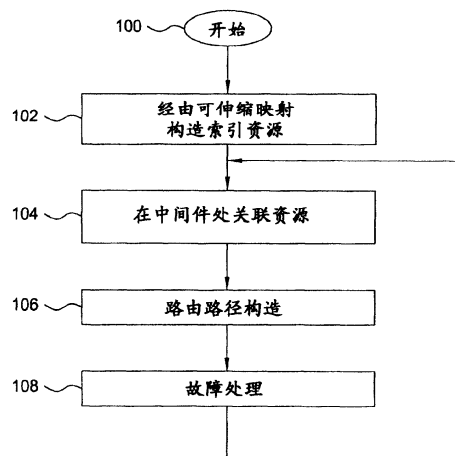
权利要求书 3 页 说明书 13 页 附图 7 页

[54] 发明名称

虚拟化网络资源的方法和装置

[57] 摘要

一种用于为合作应用管理通信的方法和装置。中间件使用网络，应用和中间件资源信息以及用户信息一起，经由实现通信覆盖树的分层控制结构来便于应用资源与用户之间的通信。该覆盖树考虑到网络约束，并且基于用户和基于他们的通信约束来实现。中间件提供用户索引标识符，其告知应用，特定用户具有在应用空间特定部分中的通信兴趣。为了将数据发送到在应用空间特定部分中感兴趣的用户，应用发送该数据以及用户索引标识符列表。然后中间件将用户索引标识符与个体用户相关联，并且沿着覆盖树将发送的数据路由至个体用户。



1. 一种通信网络, 包括:

具有网络约束的多个网络资源, 网络资源包括由具有应用空间的应用控制的应用服务器; 以及

连接到包括多个应用服务器的所述多个网络资源的中间件服务器, 所述中间件服务器用于基于包括所述网络约束和关于所述应用服务器的属性信息建立属性空间, 该中间件服务器用于索引多个应用服务器以将它们的位置反映在所述属性空间中, 该中间件服务器还用于基于网络约束和基于应用空间实现多个应用服务器与中间件服务器之间的通信覆盖树, 其中, 网络约束的改变导致所述中间件服务器实现新的通信覆盖树。

2. 根据权利要求 1 的通信网络, 还包括多个用户, 每个具有在所述应用空间中的通信兴趣, 并且其中中间件服务器根据通信兴趣索引多个用户以形成用户索引标识符; 以及其中通信覆盖树基于多个用户和基于他们的通信兴趣来实现。

3. 根据权利要求 1 的网络, 还包括多个用户, 每个具有在所述应用空间中的通信兴趣, 并且其中中间件服务器根据通信兴趣索引多个用户以形成用户索引标识符; 其中中间件服务器索引所述多个用户的网络地址。

4. 根据权利要求 2 的网络, 其中中间件服务器经由应用服务器向所述应用提供用户索引标识符, 其目的是用于告知应用, 由所述索引标识符标识的特定用户具有在应用空间特定部分中的通信兴趣。

5. 根据权利要求 4 的网络, 其中应用服务器向所述中间件服务器发送用户列表和将分发到所述用户列表上的用户的数据, 使得所述中间件服务器识别所述用户列表上的所述用户的网络位置, 并且使得所述中间件服务器发送将分发到所述用户列表上的用户的网络地址的数据。

6. 根据权利要求 5 的网络, 其中所述应用使用所述用户索引标识符来产生当在所述应用空间中通知事件发生时将被通知的用户的列表,

并且使得所述用户列表在通知事件发生时被发送。

7. 根据权利要求 5 的网络，其中所述应用服务器使用应用程序接口来发送所述用户列表和所述数据。

8. 一种虚拟化网络资源以支持在包括应用服务器和具有通信兴趣的用户的网络中的合作通信的方法，该方法包括步骤：

构造可伸缩网络映射；

根据应用服务器在网络中的位置索引应用服务器；

根据用户的通信兴趣索引用户；

基于应用的索引、基于用户的索引、和基于网络映射来产生通信覆盖树；以及

支持应用服务器和用户之间在通信覆盖树上的通信，其中当网络改变时产生通信覆盖树被重复。

9. 根据权利要求 8 的方法，其中网络映射还支持基于服务等级协议。

10. 根据权利要求 8 的方法，其中支持通信包括根据中间件软件来操作。

11. 根据权利要求 8 的方法，其中索引用户包括索引到网络的新用户。

12. 根据权利要求 8 的方法，其中如果应用服务器进入网络，则应用服务器被索引。

13. 一种操作通信网络的方法，包括步骤：

识别多个网络资源以及它们的网络约束；

识别由具有应用空间的应用控制的多个应用服务器；

识别多个用户和在每个用户的应用空间中的通信兴趣；以及

索引应用服务器以将它们的位置反映在属性空间中；

根据通信兴趣索引所述多个用户；

为所述多个用户的每个用户形成用户索引标识符，以便告知应用，由所述索引标识符标识的特定用户具有在应用空间特定部分中的通信兴趣；以及

基于识别的网络约束和基于索引的多个用户建立应用服务器与用户之间的通信覆盖树；其中当网络改变时建立应用服务器与用户之间的通信覆盖树被重复。

14. 根据权利要求 13 的方法，还包括索引所述多个用户的每个用户的网络位置。

15. 根据权利要求 14 的方法，还包括经由应用服务器向应用提供每个用户的用户索引标识符。

16. 根据权利要求 15 的方法，还包括基于所述多个用户的至少一个用户的通信兴趣和基于该至少一个用户的用户索引标识符，将数据从应用服务器发送到该至少一个用户。

17. 根据权利要求 13 的方法，其中多个用户的索引包括索引到网络的新用户。

18. 根据权利要求 13 的方法，其中建立通信覆盖树至少部分基于往返传播时间。

虚拟化网络资源的方法和装置

技术领域

本发明的实施方案一般地涉及网络通信。尤其，本发明的实施方案涉及使用中间件来管理应用服务器和用户之间的通信。

背景技术

大规模合作系统，例如网格计算，代表分布式计算系统中成长的一类，其使得用户能够动态且虚拟地共享应用、数据，和计算资源。合作系统的本质是计算的虚拟化，以使得用户能够以低成本获得高水平的计算能力。利用分布式计算系统的应用已经并且当前正在由学术和商业组织开发。

许多分布式计算应用服从具有服务质量（QoS）需求的服务等级协议（SLA）。因此，服从这种 SLA 协议的应用实现资源分配策略以满足 QoS 约束是必需的。例如，可能期望路由通信通过依据往返时间最接近个体用户的网络服务器。也可能期望基于用户类别和应用类型提供不同的网络服务器分配策略。满足这些约束可能需要复杂的通信网络和有效的通信管理。

通信管理技术的当前状态通常基于实现专用网络层机制。这种网络层机制经常是实现和维护起来困难的、成本高的，以及耗时的。这不管在许多专用网络层机制的通信原语中存在许多类似的事实。即使当专用网络层机制实现时，大部分都不考虑可能并且确实影响通信的实际网络条件。

因此，需要一种通信解决方案，其管理通信需求同时顾及资源条件。考虑实际网络条件的通信解决方案将特别有用。将应用从明确处理所有通信需求的任务中解除出来的通信解决方案将特别有用。

发明内容

本发明提供一种通信网络，包括：具有网络约束的多个网络资源，网络资源包括由具有应用空间的应用控制的应用服务器；以及连接到包括多个应用服务器的所述多个网络资源的中间件服务器，所述中间件服务器用于基于包括所述网络约束和关于所述应用服务器的属性信息建立属性空间，该中间件服务器用于索引多个应用服务器以将它们的位置反映在所述属性空间中，该中间件服务器还用于基于网络约束和基于应用空间实现多个应用服务器与中间件服务器之间的通信覆盖树，其中，网络约束的改变导致所述中间件服务器实现新的通信覆盖树。

本发明还提供一种虚拟化网络资源以支持在包括应用服务器和具有通信兴趣的用户的网络中的合作通信的方法，该方法包括步骤：构造可伸缩网络映射；根据应用服务器在网络中的位置索引应用服务器；根据用户的通信兴趣索引用户；基于应用的索引、基于用户的索引、和基于网络映射来产生通信覆盖树；以及支持应用服务器和用户之间在通信覆盖树上的通信，其中当网络改变时产生通信覆盖树被重复。

本发明又提供一种操作通信网络的方法，包括步骤：识别多个网络资源以及它们的网络约束；识别由具有应用空间的应用控制的多个应用服务器；识别多个用户和在每个用户的应用空间中的通信兴趣；以及索引应用服务器以将它们的位置反映在属性空间中；根据通信兴趣索引所述多个用户；为所述多个用户的每个用户形成用户索引标识符，以便告知应用，由所述索引标识符标识的特定用户具有在应用空间特定部分中的通信兴趣；以及基于识别的网络约束和基于索引的多个用户建立应用服务器与用户之间的通信覆盖树；其中当网络改变时建立应用服务器与用户之间的通信覆盖树被重复。

在一种实施方案中，本发明的原理一般规定中间件通信解决方案，其为合作应用管理通信，使得应用从明确处理通信的任务中解除出来。这种中间件通信解决方案使用多个属性，例如网络条件，应用逻辑，以及应用服务器资源，来建立网络通信并处理应用与用户之间服务等级协议（SLA）需求中的非均匀性。

本发明的实施方案获得关于特征例如中间件资源；网络资源例如应用服务器，网络服务器，节点容量，往返传播时间以及其他网络因素；以及用户和他们的通信约束（例如 SLA/QoS 需求）的属性信息。基于属性信息，中间件索引应用资源（服务器）以将它们的位置反映在属性空间中。当用户接触中间件时，中间件索引那些用户以将他们的个体兴趣反映在匹配他们需求的应用空间部分中。然后，中间件经由实现应用资源和个体用户之间的通信覆盖树的分层控制结构来便于应用资源（服务器）和用户之间的通信。该覆盖树考虑到关于用户和关于他们的通信约束的网络约束，例如网络节点转送容量和往返传播时间而实现。中间件将用户索引标识符经由使用通信覆盖树选择的应用服务器提供到应用。那些标识符告知应用，特定的用户（仅由索引标识符标识的）具有在应用空间特定部分中的通信兴趣。当应用想要发送数据到应用空间特定部分中感兴趣的用户时，应用经由提供有用户索引标识符，和具有在应用空间该部分中的兴趣的用户的用户索引标识符列表的应用资源来发送该数据。中间件将用户索引标识符与个体用户相关联，并且将发送的数据沿着通信覆盖树路由至个体用户。

在本发明的实施方案中，如果影响用户的属性的显著改变发生，那么该用户重新映射到通信覆盖树中。

本发明的实施方案利用应用程序接口（API）来与中间件通信。

附图说明

所以，本发明的上述特征可以详细理解的方式，上面简要概述的本发明的更特别描述可以参考实施方案而获得，实施方案中一些在附随附

图中说明。但是，应当注意，附随附图仅说明本发明的典型实施方案，因此不被看作是其范围的限制，因为本发明可以允许其他同等有效的实施方案。

图 1 说明根据本发明的原理虚拟化通信网络的流程图；

图 2 说明通信网络及其资源；

图 3 说明分配 ID 的方法；

图 4 说明通信覆盖树；

图 5 说明创建和维护覆盖树的操作的流程图；

图 6 说明创建和修改多点传送树的方法的流程图；以及

图 7 说明中间件服务器。

为了便于理解，完全相同的参考数字已经被使用，在任何可能的情况下，以指定图公用的完全相同的元素。

具体实施方式

本发明使用虚拟化通信网络的通信中间件将信息从合作应用路由至应用用户。现在参考图 1，虚拟化过程在步骤 100 开始，并且在步骤 102 通过经由可伸缩映射构造索引资源来继续。随后依次为在中间件关联资源的步骤 104，路由路径构造的步骤 106，以及使用软状态协议处理故障的步骤 108。图 1 显示通过循环返回到步骤 104 而连续的虚拟化过程，以反映中间件处可能的资源再关联。

虚拟化过程在图 2 中说明的并且与本发明的原理一致的通信网络 200 的上下文中执行。网络包括运行在一个或多个应用服务器 202A，202B，202C，和 202D 上的应用 202。应用服务器经由网络链接 203A，203B，203C，和 203D 连接到中间件 204。中间件可以包括与用户 206A-206I 通信的中间件服务器 204A，204B，和 204C。应用服务器，中间件服务器，和用户 206A-206I 可以地区地分布。中间件服务器 204A，204B，和 204C 经由链接 207A-207I 链接到 206A-206I。另外，中间件服务器可以经由网络服务器 212 连接到用户，应用资源，其他中间件服务器，或者任何其他类型的网络节点 210。

应用服务器 (202A-202D) 运行由用户 206A-206I 使用的应用 202 的一个或多个实例 (或者多个应用)。中间件 204, 经由其中间件服务器 204A-204C, 基于随后描述的多个属性管理网络通信。中间件 204 经由建立跨越应用资源, 个体用户 206A-206I, 和网络节点 210 之间的通信覆盖树的分层控制结构, 并且通过中间件服务器 204A-204C, 来实现应用服务器 (202A-202D) 和用户 206A-206I 与网络节点 210 之间的通信。例如, 中间件可以将应用服务器 (202A-202D) 与中间件服务器 (204A-204C) 以及与特定用户 (206A-206I) 相关联, 使得该应用服务器与该用户之间通过该中间件服务器的通信具有最短的往返传播时间。在另一个实例中, 中间件服务器 204A-204C 可以基于用户类别和应用类型提供不同的应用服务器到用户的分配策略。

为了完成经由可伸缩映射构造索引资源的步骤 102, 中间件 204 提供有属性信息, 使得整个网络空间可以划分成一组区域。该网络空间包括: 中间件服务器 204A-204C (一般地, 中间件资源); 网络资源例如应用服务器 (202A-202D), 网络服务器, 节点容量, 往返传播时间和其他网络约束; 以及用户和他们的通信约束 (例如 QoS)。基于该信息, 在步骤 102, 中间件通过基于由应用和/或用户指定的度量, 例如网络延迟形成分区来关联资源。

例如, 可以期望基于分配给不同网域的 AS (自治系统) 前缀来划分网络空间。在这种索引方案中, 所有 AS 前缀被收集, 然后期望的箱室 (bin) 集合被创建, 每个箱室被分配前缀集合的一个分区。基于橱柜前缀匹配, 箱室 j 可以被分配前缀 192.X.X.X 和 200.X.X.X 的全部。接下来, 每个箱室被分配代表的服务器/控制器 - 该服务器/控制器可能是物理的或者逻辑的。该分配基于期望度量, 比方说, 前缀集合到服务器 IP 地址的橱柜匹配来执行。如果服务器 (i) 分配给箱室 (j), 分配可能是该服务器 (i) 将管理由箱室 (j) 定义的网络空间分区。总之, 每个箱室将被分配服务器的索引空间, 以为它管理的用户分配反映虚拟空间的该关联的 ID。应当理解, 该划分实例仅仅是建议的实例, 而不应当理解成以任何方式限制索引空间。索引可以基于经由主动和/或被

动监控收集的度量。

注意，上面的实例说明定义网络映射和将一组服务器关联到网络空间的每个部分的一种可能方法。网络映射构造不需要是被动的。例如，可能使用测量的网络延迟，然后依据网络延迟创建彼此接近的一组分区或区域来构造映射。为了便于这种基于测量的映射构造，外部测量设施，例如基于路标的测量可以使用。例如，参考 T.S.Eugene Ng 和 Hui Zhang 的“Predicting Internet Network Distance with Coordinates-Based Approaches”，INFOCOM’02, New York, NY, June 2002 (“使用基于坐标的方法预测因特网网络距离”，INFOCOM’02, 纽约, NY, 2002年6月)。

图 1 中的循环反映用于最初索引的属性的严重改变可能需要重新划分分配空间的事实。假设那将是在非常大的时标上发生的罕见事件。简单的解决办法是重新计算索引的划分和重新分配。作为实例，考虑基于 AS 前缀的固定装箱的先前说明。假设因网络条件的改变，箱室 i 中的 AS 前缀依据网络延迟而发现更接近箱室 j 中的前缀。在这种情况下，所选前缀分配给箱室 j 。因该重新分配，将存在 ID 到属于所选 AS 前缀的用户的分配的改变。该重新分配需要通知应用服务器并改变在中间件服务器中维护的关联信息。考虑例如属于箱室 (i) 的、早期由前缀 9.2.X.X 描述的子网。因网络条件的改变，稍后将发现，9.2.X.X 更接近由箱室 (j) 管理的子网，因为 9.2.X.X 变得更接近位于箱室 (j) 中的 11.X.X.X。在这种情况下，划分被调节以满足新的划分条件。因此，由 9.2.X.X 管理的空间现在将属于箱室 (j)。作为该重新聚集过程的结果，新的代表可能为该箱室而选择。一旦更新，早期由该箱室 (j) 管理的 ID 的重新分配可以被执行以反映该改变。因此，索引将改变以反映该适应过程。

图 2 说明被索引以将它们的位置反映在属性空间中的服务器 (202A-202D)。当用户 (206A-206I) 接触中间件 204 时，中间件索引那些用户以将他们的个体兴趣反映在匹配他们需求的应用空间部分中。然后，中间件 204 可以依据关于用户，关于他们的通信约束，以及关于

应用资源的网络约束，例如网路节点转送容量和往返传播时间，来实现通信覆盖树。

当用户打算使用应用服务器（202A-202D）时，他的请求重定向到中间件服务器（204A-204C）。基于他的需求和他在应用属性空间中的定位，与虚拟空间相关联的适当的簇被确定。一旦做出该选择，虚拟 ID 以这样一种方式分配给用户，即它反映其与该簇的关联。一种简单的索引是将簇 ID 作为前缀加到用户 IP 地址以得到他的虚拟 ID。然后，应用服务器会通过将用户的虚拟 ID 宣布到服务器来被告知关于它的用户。

作为实例，考虑想要依据网络延迟最接近的服务器的用户。在这种情况下，用户从由依据网络延迟最接近用户的服务器管理的 ID 空间中分配虚拟 ID。换句话说，索引从分配给由该服务器管理的虚拟空间的索引空间中分配该用户。该分配的 ID 反映用户与该服务器的关联。一种可能的方法是将服务器 ID 作为前缀加到用户的 IP 地址以为用户产生 ID。假设该属性知道索引，适当的服务器分配给该用户。适当的控制消息发动到该所选服务器以使用用户分配的虚拟 ID 通信来自用户的请求。

图 3 说明创建虚拟 ID 分配的步骤。过程在步骤 300 开始并且在步骤 302 通过将网络空间划分成与网络前缀相关联的箱室而在网络属性空间中继续：例如箱室 (i) ->9.2.10.xxx。然后，在步骤 304，节点使用前缀匹配分配给箱室。然后，在步骤 306，新的参考节点在箱室 (j) 中计算用于网络延迟定位（当节点增加和从箱室中移除时，延迟定位服务器的集合改变）。在步骤 308，节点延迟测量获得。

使用可获得的节点延迟测量，在步骤 310，在通信兴趣空间中，节点分配给与他们的通信兴趣重叠的通信组。然后，在步骤 312，映射到通信组的服务器节点被选择。然后，在步骤 314，节点虚拟 ID 基于网络箱室分配，延迟测量和所选服务器节点的 ID 来分配。然后过程在步骤 316 停止。

在中间件处关联资源的步骤 104 也包括在中间件服务器中维护索引

与每个用户实际 IP 地址的关联。保存该关联信息的中间件服务器根据管理中间件服务器或者适当用于共享管理负载的策略的集合的控制层次来确定。中间件服务器控制协议可以决定在其他中间件服务器复制该信息。用户被返回反映他应该接触以进一步在应用空间中通信的中间件服务器的描述符。做好准备以隐藏中间件服务器中真实身份 (IP 地址等) 到虚拟 ID 的映射。

构造路由路径的步骤 106 典型地采取创建通信覆盖树的形式。图 4 说明应用资源 A1-A7 和用户 U1-U7 之间经由中间件 M1-M2 和节点 N1-N6 通信的通信覆盖树 400 的实例。如所示, 节点 N1 将 A1-A3 聚集在一起并且路由通信通过 M1, 节点 N2 将 A4-A5 聚集在一起并且路由通信通过 M1 和 M2, 而节点 N3 将 A6-A7 聚集在一起并且路由通信通过 M2。另外, 节点 N4 将 U1, U4 和 N5 链接在一起。节点 N5 也连接到 M2 和连接到 U6。M2 也直接连接到 U3 和连接到节点 N6, 节点 N6 又链接 U5 和 U7。因此, A1 可以经由节点 N1 发送数据到 M1, 并且 M1 可以将该数据路由至 N4 到 U1。应当理解, 覆盖树 200 由中间件基于多个属性来配置。

虽然我们给出关于树构建的方法, 其他方法存在。参考例如 “Enabling Conferencing Applications on the Internet using an Overlay Multicast Architecture”, Yang-hua Chu, Sanjay G. Rao, Srinivasan, Seshan and Hui Zhang, Proceedings of ACM SIGCOMM, San Diego, CA, August 2001 (“使用覆盖多点传送体系结构在因特网上允许召开会议应用”, Yang-hua Chu, Sanjay G. Rao, Srinivasan, Seshan 和 Hui Zhang, ACM SIGCOMM 学报, San Diego, CA, 2001 年 8 月); “A Case For End System Multicast”, Yang-hua Chu, Sanjay G. Rao and Hui Zhang, Proceedings of ACM SIGMETRICS, Santa Clara, CA, June 2000, pp1-12 (“终端系统多点传送的一个例子”, Yang-hua Chu, Sanjay G. Rao 和 Hui Zhang, ACM SIGMETRICS 学报, Santa Clara, CA, 2000 年 6 月, pp1-12); 以及 “Scalable Application Layer Multicast,” Suman Banerjee, Bobby Bhattacharjee, Christopher

Kommareddy, ACM Sigcomm, August 2002 (“可伸缩应用层多点传送”，Suman Banerjee, Bobby Bhattacharjee, Christopher Kommareddy, ACM Sigcomm, 2002年8月)。

将被告知任何事件的用户的集合依赖于应用逻辑和会话管理模块。应用使用应用程序编程接口(API)来与中间件服务器通信或者通过会话管理模块。给定通知事件和一组用户,中间件服务器确定必须接触的用户真实身份。使用该查找信息,适当的用户通过构造连接这些节点的有效数据通路覆盖来通知。可以使用的两种类型的覆盖包括不依赖于会话的公用覆盖和依赖于会话的每个会话的覆盖。虽然公用覆盖构造起来简单,依据带宽使用率和等待时间它可能不是最有效的。因此,在一些应用中,特定每个任务的覆盖路径是有用的。

图5说明使用虚拟ID和分层方法构造特定每个会话的覆盖路径的方法500。方法500在步骤502开始并且以确定节点与多少个独特区域相关联来继续。在步骤504,覆盖路径通过挑选这些区域的每个中的领导者来建立以连接到那些区域。每个区域中的领导者可能是区域中的任意节点或者它可能基于被告知的选择来确定。一旦每个区域中的领导者被挑选,在步骤506,连接领导者节点的树通过考虑网络空间中的相对距离来建立(这应当可以从建立的度量中获得)。如果这种信息不可获得,虚拟ID可以用于每个簇以确定树使得沿着每个树路径的延迟达到最小。接下来,在步骤508,每个区域中的覆盖在该簇中从每个簇领导者节点到剩余节点来构造。如果详细的信息可以获得,SLA特定覆盖被构造,使得SLA被满足。如果这种信息不可获得,树构建算法可以用来连接该簇中作为会话一部分的所有节点。

在一些条件下,可能存在适应覆盖路径的需求。首先,如果节点在会话中到达或离开,或者如果存在改变网络和通信空间的聚集/划分的网络条件的改变。在步骤510,做关于节点是否到达或离开的询问。如果是,在步骤512,做是否需要改变领导者节点的确定。如果不改变,在步骤514,对覆盖树做局部改变。但是,如果需要改变领导者节点(例如领导者节点离开),在步骤516,创建新的覆盖树。该改变传播到

上层，使得簇之间的路径同样被建立。

如果节点没有达到或离开，在步骤 518，做关于网络和/或通信兴趣的改变，簇内全体成员是否已经改变的询问。如果是，跳到步骤 512，确定领导改变是否需要。然后步骤 514 和 516 适当地运行。如果不是，循环返回以确定节点是否已经达到或离开。以这样一种方法，对网络的改变被监控并且网络被适当地校正。

构造的覆盖树适应节点网络属性和组员的改变。适应方法在图 6 中显示。该方法在步骤 602 开始并且通过确定节点网络属性是否已经改变而在步骤 604 继续。如果是，跳到步骤 606（其随后讨论）。但是，如果在步骤 604，确定是节点网络属性还没有改变，方法 600 继续到步骤 608，在那里做关于组员是否已经改变的确定。如果没有，循环返回到步骤 604。以这样一种方式，对影响构造的覆盖树的改变的连续监控被执行。如果节点网络属性已经改变（在步骤 604），或者如果组员已经改变（在步骤 608），在步骤 606，多点传送树被重新计算。然后，在步骤 610，基于计算的多点传送树，先前的多点传送树被修改。

前述在图 2 中所示的网络中执行。中间件 204 将用户索引标识符经由应用服务器 202A-202D 提供到应用 202。提供有特定用户索引标识符的应用服务器从通信覆盖树 200 中选择。用户索引标识符告知应用 202，仅由索引标识符标识的特定用户具有在应用空间特定部分中的通信兴趣。因此，应用 202 变得认识到，用户（206A-206I）具有在应用特定部分中的兴趣，并且通信已经经由特定中间件服务器（204A-204C）通过特定应用服务器（202A-202D）启动。但是，应用 202 不知道用户的身份或者该用户在通信覆盖树 200 中的位置。

基于应用的逻辑和会话信息，当应用 202 确定它应当发送数据到应用空间特定部分中感兴趣的用户时，应用 202 经由提供有用户索引标识符，以及具有在应用空间该部分中的兴趣的用户的用户索引标识符列表的应用服务器（202A-202D）来发送该数据。然后，中间件 204 将用户索引标识符与个体用户（206A-206I）相关联，并且将发送的数据沿着通信覆盖树 200 路由至个体用户（206A-206I）。

因为中间件 204 管理用户 (206A-206I) 和应用 202 之间在具有节点的网络上的通信, 以有效地运行, 通信网络 200 应当根据特定的通信协议来操作并且应当支持通信覆盖树 200。通信协议通过向应用开发者提供用于将应用 202 和用户 (206A-206I) 与中间件 204 软件相接口的应用程序编程接口 (API) 来建立。API 实现使得应用开发者能够开发应用 202 的适当且标准化的通信原语, 使得用户 206A-206I 可以通过中间件 204 连接到运行在应用服务器 202A-202D 上的应用 202。

支持通信覆盖树需要实现与树构造和树使用相关的步骤。图 3 说明支持通信覆盖树的方法 300。该方法在步骤 301 开始并且通过使用可伸缩映射索引网络资源在步骤 302 继续。该可伸缩索引标识应用服务器 (202A-202D) 和资源, 中间件服务器 (204A-204C), 网络节点 (210), 以及各种网络通信参数例如节点转送容量, 节点之间的往返时间延迟, 包损失率, 以及其他网络约束。因此, 映射代表关于覆盖树 200 将在其中操作的网络环境而收集的数据。应当理解, 步骤 302 正在进行, 其中网络环境是动态的。SLA 或者影响用户感知到 SLA 的条件任何显著改变可以启动网络资源的新的索引。

图 2 应用服务器 A1-A3 被聚集以经由节点 N1 通信。聚集可以对于应用服务器 202A-202D, 链接 203, 节点 212, 和中间件服务器 204 的给定集合而执行, 以形成虚拟空间。例如, 欧几里得空间可以用来将应用服务器划分成簇, 然后其使用矢量结构中适当数目的簇位来索引。接下来, 对于每个簇, 索引被形成以指向特定的应用服务器。这种聚集方法允许应用服务器 202A-202D 的移除和增加。注意, 用来创建该索引的实际属性可能依赖于应用的限制和需求, 应用服务器 202A-202D 的限制和需求, 链接 203 的限制和需求, 以及中间件服务器 204A-204C 的限制和需求。再次, 聚集是正在进行的过程, 其中网络中的严重改变可以启动新的聚集。严重改变是影响网络支持用户和/或应用服务器的需求的能力的任何网络改变。严重改变应当是仅在大的时标上发生的罕见事件。但是, 如果严重改变被检测到, 步骤 302 的网络资源索引可以重新确定, 继之以在步骤 303 的聚集。作为选择, 试图不扰乱最初分配

的网络资源索引的增加方法可以实现。

在网络资源在步骤 303 聚集之后，在步骤 304，用户被索引。当用户进入通信网络 200 时，用户的需求和兴趣输入到中间件服务器 204A-204C，从而输入到中间件 204。需求和兴趣可以作为用户矢量发送，其中各个位代表预先指定的特征。然后，在步骤 306，基于用户需求，用户兴趣，用户在应用属性空间中的位置，以及用户和中间件服务器之间的网络资源，中间件为用户确定适当的应用服务器簇。例如，该关联可能反映基于时间或者基于可接受的包数据丢失到用户最近的应用服务器。基于使用的索引方法，在步骤 308，用户被分配索引标识符，其经由适当的应用服务器 202A-202D 传递到应用 202，并且发送返回用户。索引标识符告知应用 202，用户 206A-206I 在应用空间特定部分中感兴趣，并且通信将通过特定的应用服务器 202A-202D 到达特定的中间件服务器 204A-204C。用户也被返回通信描述符，其识别它应当接触以在应用空间中进一步通信的中间件服务器 204A-204C。

但是，中间件 204 例如通过索引用户的 IP 地址连同用户索引标识符一起，来跟踪真实用户及其在通信覆盖树 200 中的位置。保存该关联信息的真实中间件服务器 204A-204C 根据管理中间件服务器的中间件控制层次来确定。另外，中间件服务器控制协议可以决定在其他中间件服务器的一个或多个中复制该信息。

在步骤 308 之后，在步骤 310，应用的应用逻辑和/或会话管理模块确定将基于通知事件的发生接收数据的用户 206A-206I 的列表。在步骤 312，当通知事件被检测到时，应用 202 将数据连同将接收数据的用户 206A-206I 的列表一起，发送到预先确定的中间件服务器 204A-204C。然后在步骤 314，中间件 204 确定将被接触的用户 206A-206I 的真实身份（IP 地址），然后发送数据到那些用户。实际上，应用 202 或会话管理模块使用提供的 API 来描述用户集合到适当的中间件服务器 204A-204C。

假设用户 206A-206I 与应用服务器 202A-202D 的关联在中间件服务器 204A-204C 处维护，在中间件服务器故障到来时支持不中断的通信

是有用的。步骤 108，故障处理由两组机制来获得。第一组是在除指定服务器之外的多个中间件节点处复制状态。第二组是当检测到父服务器已经失败时，舍弃跨越到备份服务器。几种方法可以用来检测故障。例如，每个用户可以使用保持活性消息以检查节点故障。在节点故障到来时，备用服务器被接触以重新建立连接。同样除了节点故障之外，如果存在性能退化，可能选择备份服务器（或者重新选择服务器）。

图 7 描绘中间件服务器 700 的实施方案的高级框图。中间件服务器 700 包括处理器 710 以及用于存储控制程序 725 和数据结构 727 的计算机可读介质 720。处理器 710 与常规支持电路 730 例如电源、时钟电路、高速缓冲存储器等以及帮助执行存储在存储器 720 中的软件例程的电路合作。同样地，应当注意，在这里作为软件过程讨论的过程步骤的一些可以在硬件中实现，例如作为与处理器 710 合作以执行各种步骤的电路。中间件服务器 700 也包含输入-输出电路 740，其形成与中间件服务器通信的各种网络组件（例如硬盘驱动器，光盘驱动器等）之间的接口。

虽然中间件服务器 700 描述成根据本发明编程以执行各种控制功能的通用计算机，本发明可以在硬件中，例如作为专用集成电路（ASIC）实现。同样地，在这里描述的过程步骤打算广泛地说明成由软件，硬件，或其组合同等地执行。而且，计算机可读介质 720 是存储可以由处理器 710 运行和/或访问的信息的计算机可读存储器。

前述导致虚拟化的网络，其中通信由中间件管理和处理，这将应用和用户从必须明确地彼此通信中解除出来，并且其中通信基于网络条件。

虽然前述针对本发明的实施方案，本发明的其他和更多实施方案可以被设计，而不背离其基本范围，并且其范围由随后的权利要求来确定。

图1

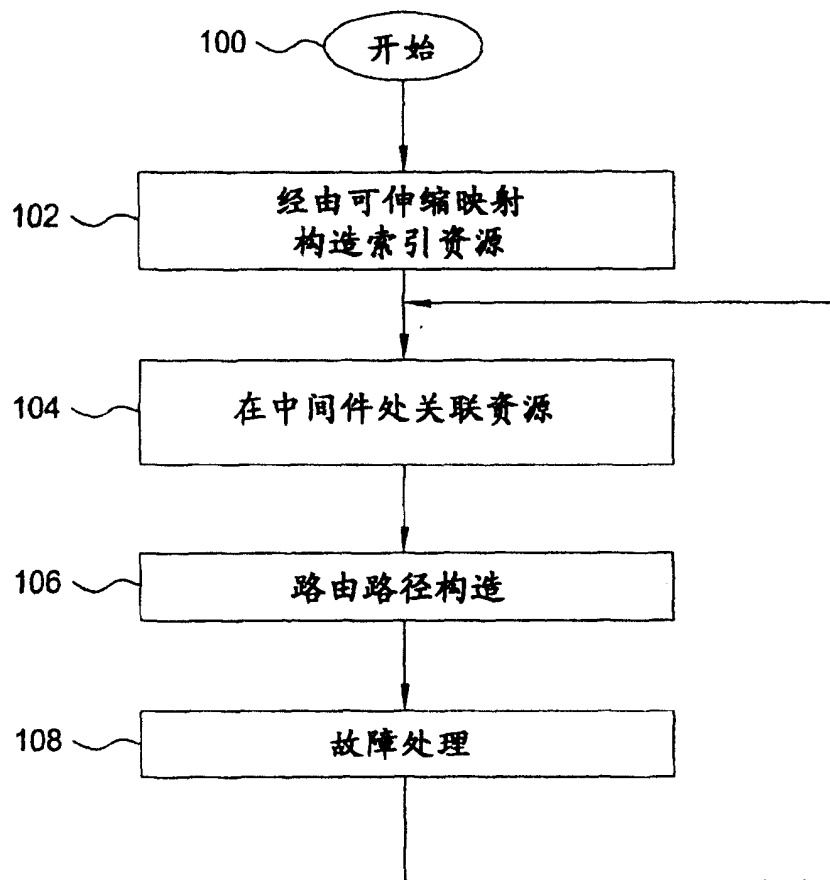


图 2

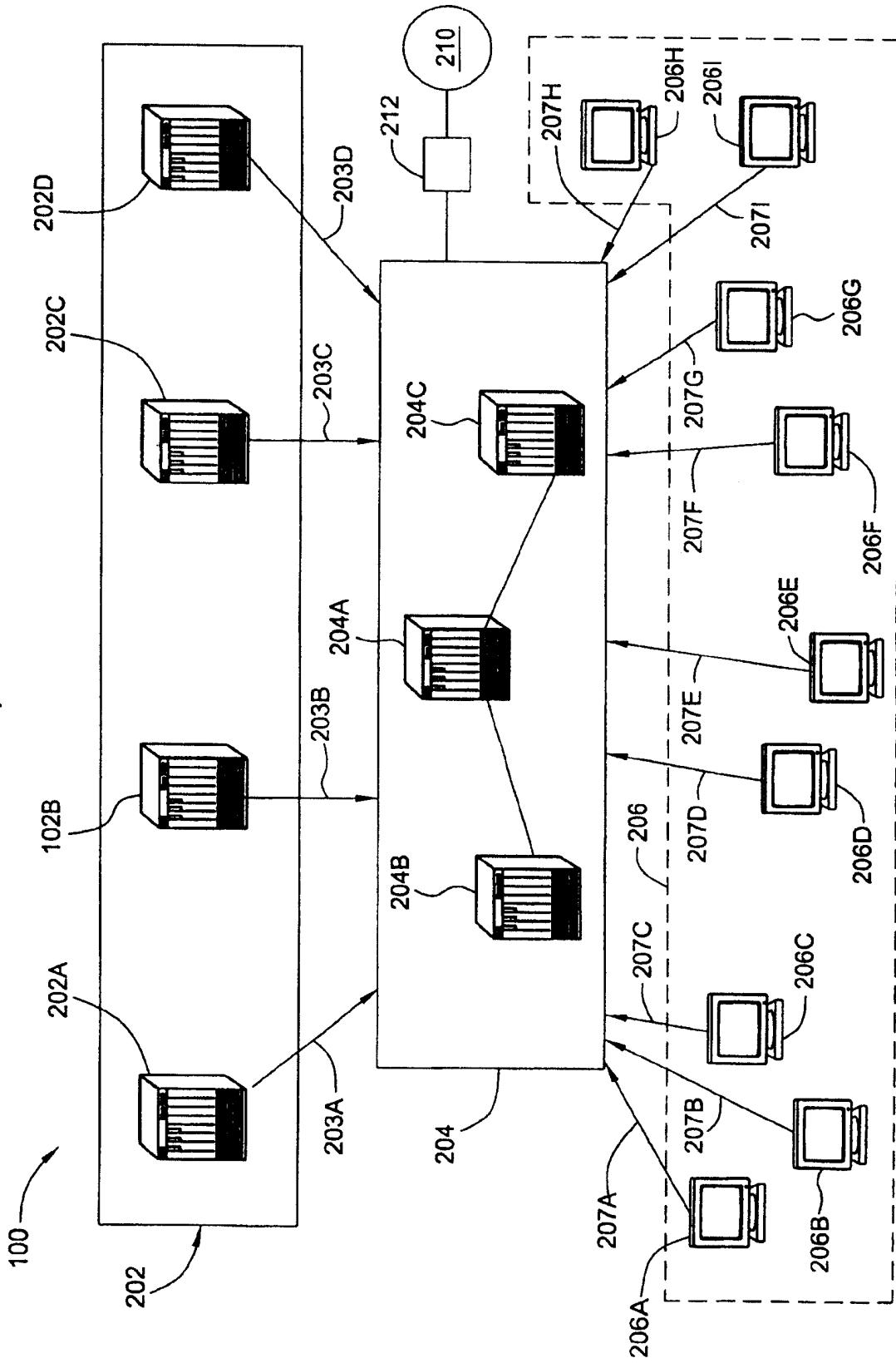


图3

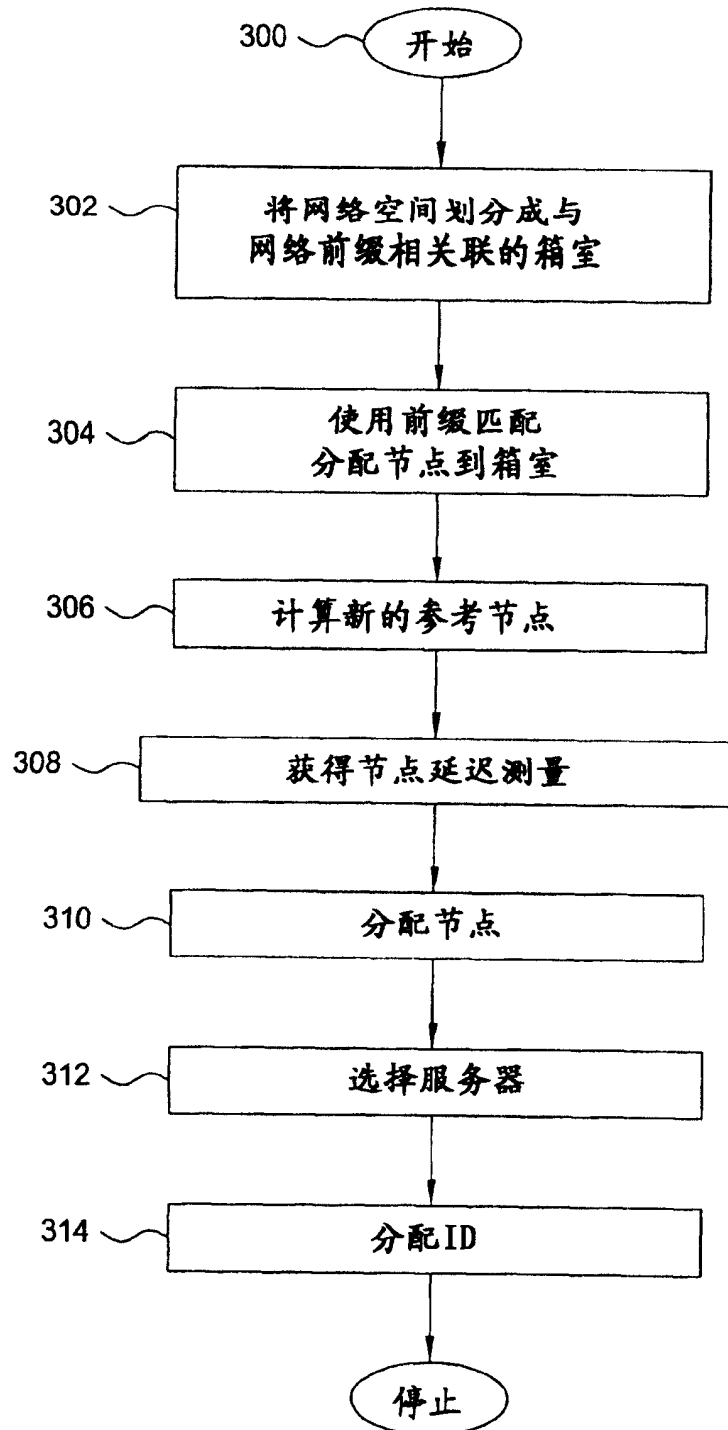


图 4

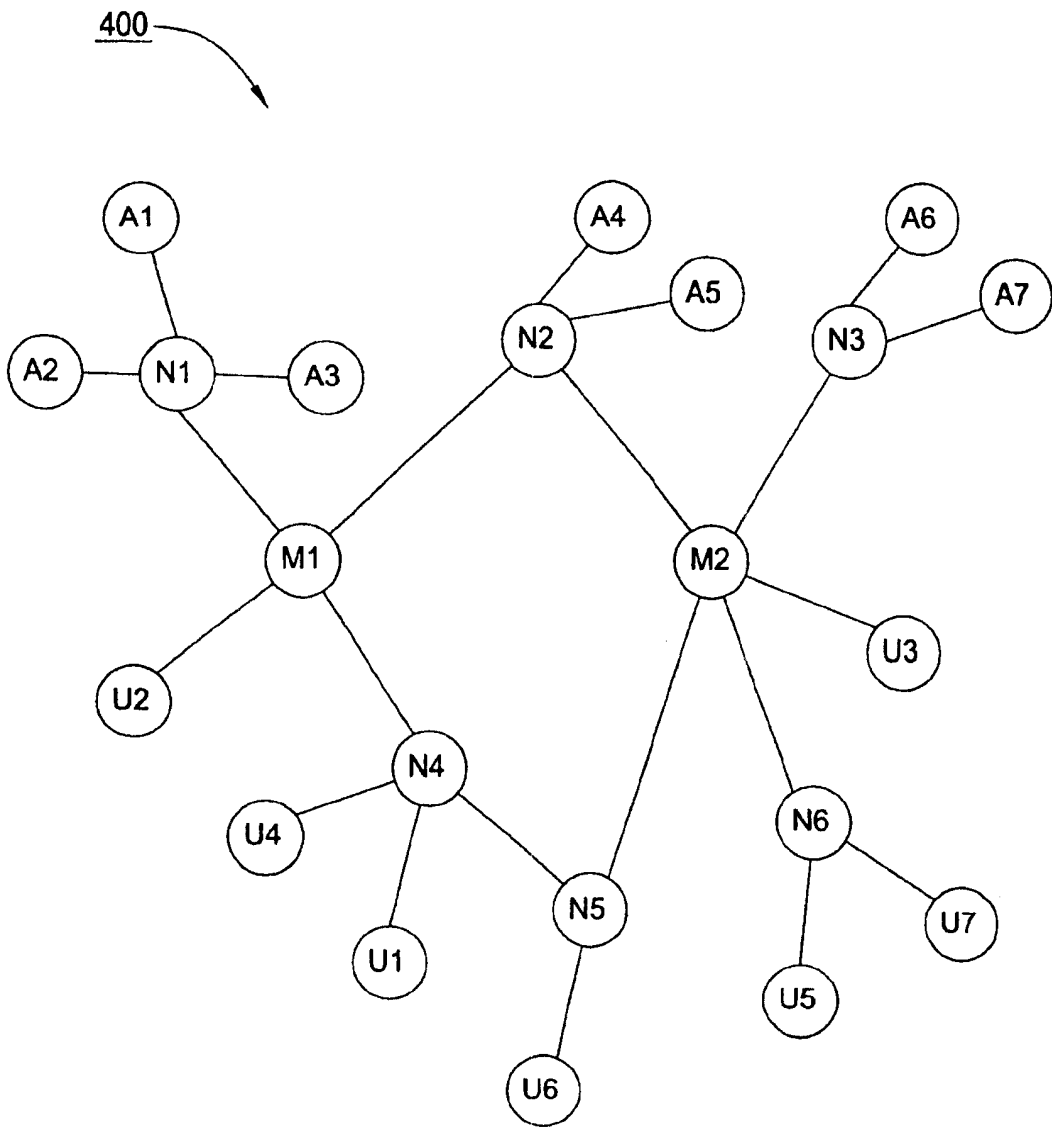


图5

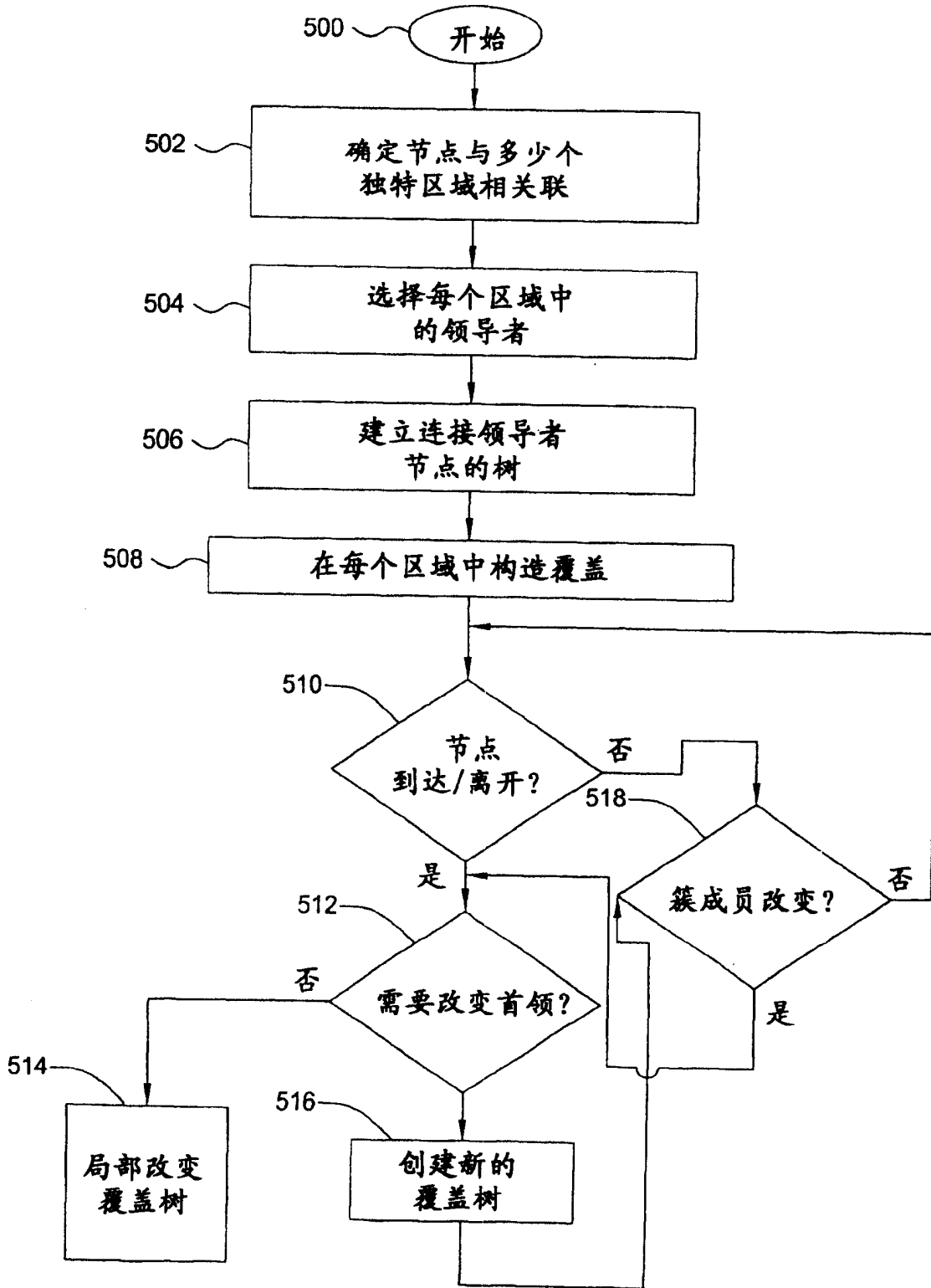


图6

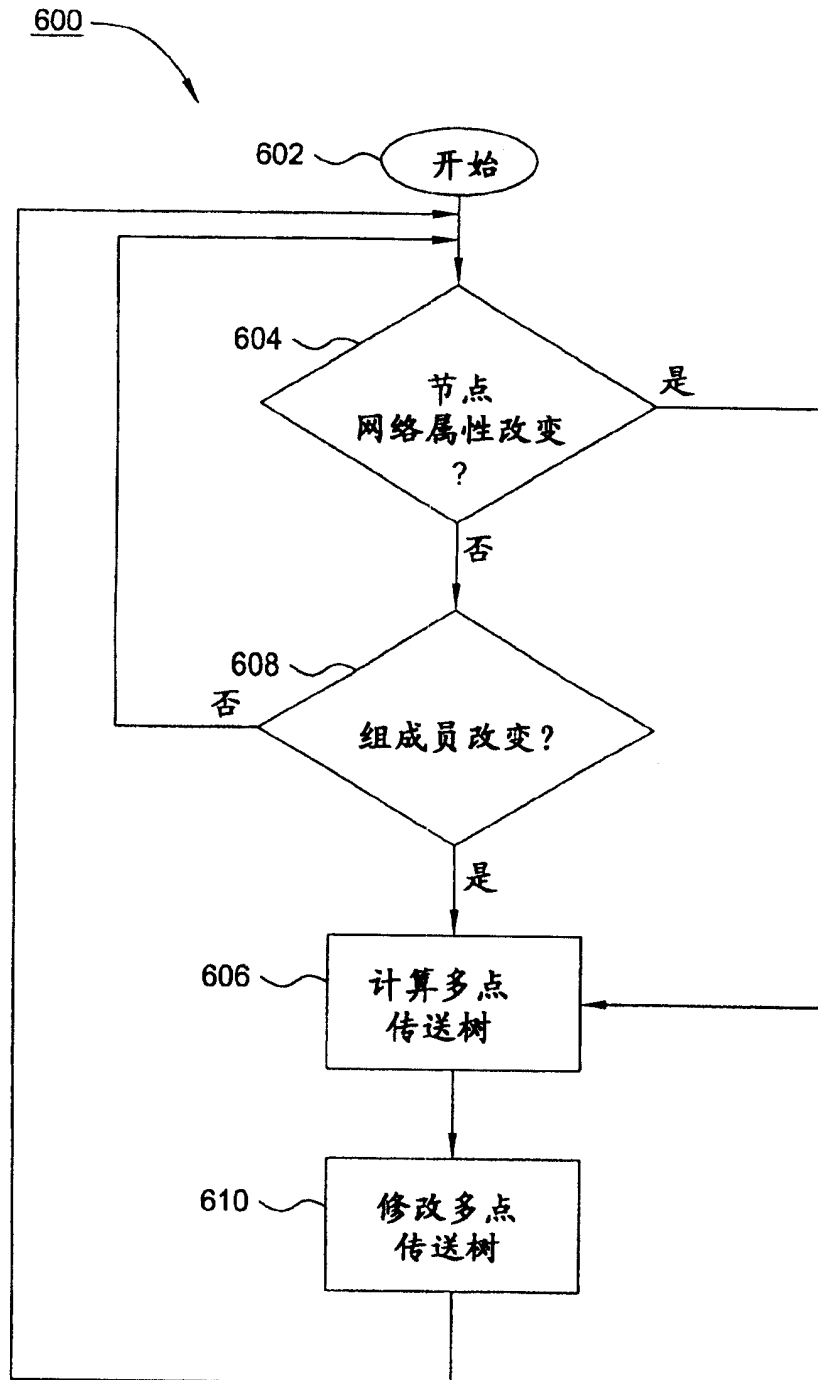


图7

